Using initialized hindcasts to assess simulations of 1970-2009 equatorial Pacific SST, zonal wind stress, and surface flux trends

Amy Solomon

CIRES/University of Colorado and PSD/ESRL/NOAA

email:amy.solomon@noaa.gov

July 3, 2014

Submitted to Journal of Climate

Abstract

Initialized decadal hindcasts are used to assess simulations of 1970-2009 equatorial Pacific SST, zonal wind stress, and surface flux trends. Initialized hindcasts are useful to assess how well the models simulate observed trends, as well as, how simulations of observed trends (due primarily to natural variability) differ from ensemble mean forecasted trends (due to the response to an increase in external forcing).

All models forecast a statistically significant warming trend in both the warm-pool and cold-tongue regions. However, while the warm-pool warming trend is within the observed estimates, the cold-tongue warming trend is an order of magnitude larger than an ENSO residual estimated using SST instrumental reconstructions. Multi-model ensemble means formed using forecasts 6-10 years from initialization with 40 ensemble members do not produce an unambiguous zonal SST gradient response to an increase in external forcing.

Systematic biases are identified in forecasts of surface fluxes. For example, in the warm-pool region all year-1 forecasts produce SST trends similar to observations but ocean mixed-layer and net surface heat flux trends with opposite sign to air-sea datasets. In addition, year-1 forecasts produce positive shortwave feedbacks on decadal time scales, while 6-10 year forecasts produce negative or statistically insignificant shortwave flux feedbacks on decadal time scales, suggesting a sensitivity to circulations forced by the initialized ocean state. In the cold-tongue region initialized ensembles forecast positive net radiative flux trends even though shortwave flux trends are negative, i.e., for increasing cloudiness. This is inconsistent with air-sea datasets, which uniformly show

32 that the net surface radiative flux feedback is a damping of the underlying SSTs.

1. Introduction

Based on estimates of centennial trends in the Indo-Pacific region from instrumental reconstructions of sea surface temperatures (SSTs) it is uncertain whether an increase in greenhouse gases has resulted in a strengthening or a weakening of equatorial Pacific zonal temperature gradients (Cane et al. 1997; Karnauskas et al. 2009; Compo and Sardeshmukh 2010; Deser et al. 2010; Meng et al. 2012; Solomon and Newman 2012; L'Heureux et al. 2013). For example, Solomon and Newman (2012) find a strengthening of the zonal SST gradients due to a weak SST cooling trend in the eastern equatorial Pacific since 1900 in four SST reconstructions that is not statistically significant relative to natural variability, while studies such as Tokinaga et al. (2012) find a weakening of the zonal SST gradient trend due to eastern equatorial Pacific SSTs warming faster than SSTs around the Maritime Continent when using a shorter period (1950-2005). These studies indicate that the observational record is not sufficient to prove or disprove opposing theories of the response of eastern equatorial Pacific SSTs to an increase in external forcing.

There is likewise some uncertainty about whether the related atmospheric Pacific Walker circulation (hereafter Walker circulation or Walker cell) has weakened or strengthened in the twentieth century (Clarke and Lebedev 1996; Vecchi et al. 2006; Power and Smith 2007; Vecchi and Soden 2007; Bunge and Clarke 2009; Karnauskas et al. 2009; Deser et al. 2010; Solomon and Newman 2012). These uncertainties are primarily due to large interannual variability, sparsity of data, and changes in the observing system in the eastern equatorial Pacific and prevent a clear verification of the response of the Walker

55 cell to external forcing in global climate model simulations of the 20th-21st centuries.

56 Natural variability and changes in the observing system also contribute to uncertainties in
57 in equatorial Pacific surface flux trends (Trenberth 2002; Wielicki et al. 2002a; Wielicki
58 et al. 2002b). However, in addition to these common uncertainties, calculations of
59 radiative fluxes at the ocean surface need to use subjective measures of cloud fraction
60 (for ship-derived estimates) or models of cloud microphysics and assumptions about
61 cloud overlap (for satellite derived estimates).

62 Due to these uncertainties in the observational record (and the potentially weak signal of
63 external forcing in the tropical Pacific, for example, see Solomon and Newman (2011)), it
64 is necessary to use coupled climate model simulations to identify the response of the
65 equatorial Pacific to an increase in greenhouse gases. Initialized hindcasts are useful for
66 this purpose since they can be used to assess how well the models simulate observed
67 trends, as well as, how simulations of observed trends (due primarily to natural
68 variability) differ from ensemble mean forecasted trends (due to the response to an
69 increase in external forcing). However, estimates of observed ENSO-residual trends
70 (trends with ENSO variability removed) need to be used to verify simulated trends in
71 initialized hindcasts for lead-times greater than approximately six years, since after this
72 time the hindcasts loose memory of the initial conditions and an ensemble mean with
73 sufficient members averages out natural variability (Branstator and Teng 2010).

74 In this study we use decadal hindcast ensembles archived in the Climate Model
75 Intercomparison Project phase 5 database (CMIP5; Taylor et al. 2012). These ensembles

are initialized yearly from 1970-2009 and run 10 times with perturbed initial conditions for each start date. The large number of simulations in this ensemble allows for robust estimates of the response of each model to an increase in external forcing. We focus our analysis on assessing simulations of equatorial Pacific SST, zonal wind stress, and surface flux trends.

We use two indices to describe, to first order, the trend pattern in the equatorial Pacific Ocean; the warm-pool index (5°N-5°S, 120°-165°E) and the cold-tongue index (5°N-5°S, 180°-70°W). The Pacific zonal mean zonal wind stress index (the τ_x index; 5°N-5°S, 120°E-70°W) is used to describe the equatorial wind stress trends. These three indices are used to describe trends in the strength of the Walker circulation, the asymmetric circulation in the tropical Pacific where convection in the warm-pool region forces upward motion that subsides in the central-eastern Pacific. In addition, we use these indices to assess simulations of surface flux trends in the warm-pool and cold-tongue regions.

We demonstrate that the decadal hindcasts do not produce an unambiguous weakening or strengthening of the Walker circulation over the 1970-2009 period, even for multi-model ensemble means of 6-10 year forecasts (a mean over 40 ensemble members beyond the influence of the initial observed state). It is argued that, even though models produce statistically significant warming trends in both the warm-pool and cold-tongue regions, the systematic response of the SST zonal gradient to an increase in external forcing over the 1970-2009 period is too weak to be identified with a 40-member ensemble.

97 Surface flux trends are assessed to identify potential biases in simulations of natural
98 variability over the 1970-2009 period, and to compare fluxes due to natural variability
99 with the response to external forcing. This analysis demonstrates that all initialized
100 ensembles used in this paper forecast year-1 warm-pool SST trends similar to
101 observations but ocean mixed-layer and net surface heat flux trends with opposite sign to
102 air-sea datasets. In addition, all 6-10 year forecasts in the cold-tongue region produce
103 positive net radiative flux trends even though shortwave flux trends are negative, i.e., for
104 increasing cloudiness. This result is inconsistent with air-sea datasets, which uniformly
105 show that the net surface radiative flux feedback is a damping of the underlying SSTs.

106 ## 2. Methods, Models and Data

107 We apply our analysis to four sets of initialized decadal hindcasts archived in the CMIP5
108 database that are initialized yearly from 1960-2009. Each start date has 10 ensemble
109 members with perturbed initial conditions. These hindcasts take into account changes in
110 external forcings such as greenhouse gases, solar activity, stratospheric aerosols
111 associated with volcanic eruptions and anthropogenic aerosols. The first two ensembles
112 use the UK Met Office coupled climate model HadCM3 configured with a horizontal
113 resolution of 2.5^{o}x2.5^{o} in the atmosphere and 1.25^{o} in the ocean (Gordon et al. 2000). The
114 HadCM3-i2 ensemble is anomaly initialized (observed anomalies and the model's mean
115 climate are used as initial conditions) and the HadCM3-i3 ensemble is initialized with
116 full fields (observed anomalies and climate mean states are used as initial conditions).
117 The third ensemble uses the Canadian Centre for Climate Modelling and Analysis
118 CanCM4 (Arora et al. 2011; Merryfield et al. 2013), and is full-field initialized. The

119 fourth ensemble uses the NOAA Geophysical Fluid Dynamics Laboratory CM2.1
120 (Delworth et al. 2006; Chang et al. 2013) and is anomaly initialized. A multi-model
121 ensemble mean (MME) is formed by averaging time series from the 4 ensembles (3
122 ensembles in Sections 3.3 and 3.4). All fields are interpolated to the HadCM3 2.5^{o}x2.5^{o}
123 grid. Only results using annual and ensemble means are presented in this study.

124 Annual mean anomalies are bias corrected as a function of lead-time, where the model
125 forecast anomaly is calculated as $Y'_{j\tau} = Y_{j\tau} - \overline{Y}_{\tau}$, where $\overline{Y}_{\tau}$ is the ensemble-average
126 forecast as a function of lead-time τ, $Y'_{j\tau}$ is the anomaly of the raw forecast with respect
127 to the ensemble average, j is the starting year. $\overline{Y}_{\tau}$ is calculated as $\frac{1}{n}\sum_{j=1}^{n} Y_{j\tau}$. However, it
128 is important to note that the trends calculated as a function of lead-time are independent
129 of the bias correction since all years used to estimate the trend have the same correction.

130 Linear trends are calculated using the method of least squares linear regression.
131 Confidence intervals are estimated using a two-tailed Student's t distribution (Bendat and
132 Piersol, 2000). Trends are estimated to be significantly different from a zero trend when
133 they exceed the 95% level.

134 Forecasted trends are calculated as follows: trends at lead year 1 are calculated from the
135 time series of the first year of the 50 start dates, trends at lead year 2 are calculated from
136 the time series of the second year of the 50 start dates, etc. In tropical Pacific SSTs, the
137 externally forced signal emerges from the signal due to initial conditions after
138 approximately 4 years and all information from the initialization is lost in approximately

139 6 years (Branstator and Teng 2010). We therefore use forecasted trends for lead years 6-
140 10 as estimates of uninitialized trends.

141 Three data assimilations and four SST reconstructions are used to verify simulated SSTs
142 and wind stress. The data assimilations are the European Centre for Medium-range
143 Weather Forecasts Ocean Reanalysis System 4 (ORAS4 or "O", Balmaseda et al. 2013),
144 the Geophysical Fluid Dynamics Laboratory Ensemble Coupled Data Assimilation V3.1
145 (ECDA or "G", Chang et al. 2013), and the Simple Ocean Data Assimilation version
146 2.1.6 (SODA or "S", Carton and Giese 2008). The SST reconstructions are the Hadley
147 Centre Sea Ice and SST dataset version 1.1 (Rayner et al. 2003), the National Oceanic
148 and Atmospheric Administration Extended Reconstruction SST version 3b dataset (Smith
149 et al. 2008), Lamont Doherty Earth Observatory SST version 2 (Kaplan et al. 1998),
150 Centennial in Situ Observation Based Estimates of SST (Ishii et al. 2005). The four
151 reconstructions are averaged to form a "best estimate", marked with an "A" in Fig. 1.
152 Also shown in Fig. 1 is the best estimate of the SST trends where ENSO variability has
153 been removed following the technique developed in Solomon and Newman (2012),
154 marked with "R".

155 Three air-sea flux datasets are used to validate the modeled surface fluxes; 1983-2006
156 Coordinated Ocean-ice Reference Experiments dataset Version 2 (CORE.2 or "C2",
157 Yeager and Large 2008), 1984-2009 Woods Hole Oceanographic Institute Objectively
158 Analyzed air-sea Fluxes (OAFlux or "O", Yu and Weller 2007), and 1983-2009 National
159 Oceanography Centre Version 2.0 Surface Flux and Meteorological dataset (NOCv2 or
160 "N", Berry and Kent 2011). Surface radiative fluxes are satellite-derived in all datasets

except NOCv2, which uses ICOADS (Woodruff et al. 1998; Worley et al. 2005) ship-based estimates of cloud cover and empirical formulas.

3. Results

3.1 SST and τ_x trends

Ensemble mean 1970-2009 warm-pool, cold-tongue, and τ_x trends as a function of lead year are presented in Fig. 1. Looking at the observed trends first (leftmost column), there is a systematic warming trend in all of the observed warm-pool time series even though these trends differ by up to a factor of 2. This is not the case for the τ_x and cold-tongue trends, where large variability obscures identifying whether these trends have increased or decreased over the 1970-2009 period. There is a tendency for a warming trend in the cold-tongue region, however only one of these time series is significant beyond the 95% level. Since large uncertainty in the cold-tongue region is primarily due to ENSO variability we also show an estimate of an ENSO residual trend (marked "R" in Fig. 1, where ENSO variability has been removed following the technique of Solomon and Newman (2012)). Removing ENSO variability reduces the cold-tongue warming trend by over 68%, which is useful in assessing the magnitude of the ensemble mean simulated trends at long leads but uncertainty in the ENSO residual index is too large to validate the sign of the simulated trends.

To identify the impact of the initialization method on the long-lead trends we use two ensembles of HadCM3 initialized hindcasts, one that employed anomaly initialization (HadCM3-i2) and one that employed full-field initialization (HadCM3-i3). Even though

182 full-field initialized hindcasts have large model drift from the observed climatology to the
183 model climatology (see Kim et al. 2012), the bias-corrected trends shown in Fig. 1 are
184 relatively insensitive to the initialization method.

185 Looking at the multi-model ensemble mean (MME) trends in Fig. 1, a significant
186 warming trend is seen in both the warm-pool and cold-tongue trends by lead year 4. This
187 is not the case for the MME τ_x trends, where even though confidence intervals narrow for
188 increasing lead years, trends for lead years 4-10 are not significantly different from zero.
189 Warm-pool trends are within the observed estimates. However, cold-tongue trends for
190 lead years 4-10 are an order of magnitude larger than the ENSO residual estimate, which
191 shows a weak statistically insignificant cooling trend.

192 The loss of information in the initial conditions in all three MME indices is seen in the
193 narrowing of the confidence intervals from lead year 1 to lead year 4, thereafter
194 confidence intervals are approximately constant. It is between leads years 1 and 4 that the
195 divergence from the initial observed ocean state estimate can be seen, for example in the
196 discontinuous jump in all MME trends between leads years 2 and 4.

197 **3.2 Relationship between the Walker cell and zonal SST gradient**

198 The scatter between 1970-2009 warm-pool minus cold-tongue SST (near-equatorial zonal
199 SST gradient) and τ_x trends from Fig. 1 is shown in Fig. 2. First, it is seen that the two
200 observational trends plotted do not produce a consistent relationship between the SST
201 gradient and τ_x trends. However, all year-1 forecasts of the trends have confidence
202 intervals that overlap the observed confidence intervals in both the SST gradient and τ_x

directions, indicating that the modeled spread is within the observed estimates. Also, it is interesting to see that the multi-model ensemble mean year 6-10 forecasts are scattered across the zero lines. Therefore, even ensembles for forecasts 6-10 years from initialization with 40 ensemble members do not produce an unambiguous response of the Walker circulation to a change in external forcing for the 1970-2009 period.

The systematic relationship between the zonal wind stress and SST gradient in the hindcasts indicates that the strength of the Walker cell is a function of the *relative* warming between the warm-pool and cold-tongue regions, as suggested by Bjerknes (1969). Therefore, even though there has been a systematic warming in the warm-pool region since 1900 (i.e., Fig. 3, Solomon and Newman 2012), the strength of the Walker cell depends on the relative warming in the cold-tongue region.

3.3 Warm-pool and cold-tongue surface energy flux trends

The net surface energy flux into the ocean is

$$Q_{net} = Q_S + Q_L + Q_H + Q_E \quad (1)$$

where Q_S is the net solar radiation, Q_L is the net longwave radiation, Q_H is the sensible heat flux, and Q_E is the latent heat flux.

a. Warm-pool fluxes

Fig. 3 shows surface energy flux trends from the CanCM4, HadCM3-i2, and GFDL CM2.1 ensembles compared to CORE.2, OAFlux, and NOCv2 fluxes for the 1983-2006

222 period (the common period between the observational estimates and models). The
223 observed fluxes indicate that the net surface energy flux trend is negative in the warm-
224 pool region (Fig. 3a). A decreasing net surface energy flux and net surface shortwave flux
225 after the 1997-1998 El Niño in the warm-pool region is consistent across the air-sea
226 datasets (results not shown). The NOCv2 ship-based air-sea flux dataset that extends
227 back to 1954 indicates that there has been a long-term decreasing cloudiness trend in the
228 warm-pool region, consistent with the findings of Norris (2005), Deser et al. (2010), and
229 the recent discussion by Bellomo et al. (2014). However, the three datasets show that
230 there has been decadal variability in cloudiness in the warm-pool region since 1983 with
231 a decrease in cloudiness until 1997-1998 and an increase thereafter. The sign of the net
232 surface longwave flux trend differs across the three datasets. Also, note the large range in
233 the latent heating trends across the three datasets.

234 The SST tendency equation can be written as

235 $$C\frac{\partial T'}{\partial t} + D_o + Q_{net} \qquad (2)$$

236 where T' is the SST anomaly, $C = c_p \rho_o H$ is the heat capacity of the ocean mixed layer,
237 is c_p the specific heat at constant pressure, ρ_o is the density of seawater, H is the mixed
238 layer depth, Q_{net} is the net surface flux into the ocean, and D_o is the mixed layer ocean
239 heat transport. For annual means, the heat storage term is an order of magnitude smaller
240 than the net surface flux and ocean heat transport terms and

241 $$D_o \approx -Q_{net} \qquad (3)$$

Therefore, the small warm-pool SST tendency trend relative to the net surface flux trend in the observational estimates requires a warming due to ocean mixed layer heat fluxes to compensate for the cooling due to net surface fluxes.

All three models produce interannual ENSO variability with a negative shortwave flux feedback, i.e., shortwave (and net) fluxes damp the underlying SSTs. However, superimposed on the interannual variability is a positive trend for both SSTs and shortwave fluxes in all three models, with year-1 net surface flux trends between 5-15 times smaller than the observational estimates. This results in all three models producing net surface and ocean mixed-layer heat flux trends opposite to the observational estimates.

For year-1 forecasts the models produce a positive shortwave feedback on decadal time scales and a negative shortwave feedback on interannual time scales. By contrast, the 6-10 year forecasts produce a negative or statistically insignificant shortwave flux feedback on decadal time scales with a compensating positive longwave flux trend. It is unclear from this analysis what causes the positive shortwave flux feedback in the year-1 forecasts in all three models, but given that this result is not found in the year 6-10 forecasts, this suggests that the positive shortwave feedback in the warm-pool region is due to atmosphere and ocean circulations forced by the initialized ocean state.

b. Cold-tongue fluxes

Cold-tongue shortwave fluxes estimated from all three flux datasets are negatively correlated with the observed interannual warm-pool fluxes discussed above, consistent

263 with an eastward shift in cloudiness during El Niño events. However, the long-term
264 increase in cloudiness seen in longer records (Norris 2005; Deser et al. 2010; Bellomo et
265 al. 2014) is not produced using the 1983-2006 satellite record and the CORE.2 and
266 OAFlux datasets produce a positive shortwave flux trend (decreasing cloudiness), while
267 the ship-derived NOCv2 dataset produces a negative shortwave flux trend. All three
268 datasets produce a negative net surface flux trend over the 1983-2006 period. Estimates
269 of net surface longwave fluxes differ across the air-sea datasets, even for CORE.2 and
270 OAFlux that use the same ISCCP-FD radiative fluxes.

271 As was seen in the warm-pool fluxes, the initialized hindcasts produce year-1 forecasts
272 with negative net surface flux feedbacks in the cold-tongue region that damp SST
273 anomalies that evolve during an ENSO event, similar to the air-sea datasets (results not
274 shown). Again, as was seen in the warm-pool region, superimposed upon this interannual
275 variability all models produce positive net surface flux trends, due to the dominance of
276 positive longwave flux trends over shortwave fluxes trends.

277 The response to external forcing (as estimated by the year 6-10 forecast ensemble mean)
278 is consistent across the 3 models and with the year-1 forecasts (except for the HadCM3-i2
279 positive year-1 shortwave flux trend). A decreasing shortwave flux trend indicates that
280 cloudiness increases as SSTs increase in the cold-tongue region. This cooling trend is
281 compensated by a positive downward longwave flux trend, resulting in positive net
282 radiative and net surface flux trends in all three models. All three ensembles produce year
283 6-10 forecasts with net surface flux trends opposite to the observational estimates. In
284 addition, latent heating trends damp the underlying SSTs but are weaker than the

observational estimates.

3.4 Cold-tongue surface energy flux feedbacks

Large uncertainty estimates and the differences across air-sea datasets in the cold-tongue region obscure relationships between surface fluxes and SSTs (Fig. 3). More systematic relationships can be identified by calculating surface energy flux feedback coefficients (units of W m^{-2} K^{-1}), which are calculated by regressing terms in the surface energy budget to the cold-tongue SSTs. The surface energy flux feedback coefficients are listed in Table 1 for the observational estimates and in Table 2 for the year 6-10 forecasts.

In all three air-sea datasets, shortwave fluxes are a negative feedback. The longwave fluxes differ across the datasets (as noted above) but the net radiative flux feedback is uniformly negative across the three observational estimates with an average value of -5.9±1.9 W m^{-2} K^{-1}. Including the turbulent heat fluxes results in an average net surface flux feedback equal to -13.1±0.9 W m^{-2} K^{-1}.

The year 6-10 forecasts also produce a negative net shortwave flux feedback but the average feedback coefficient is a factor of 4.7 smaller than the observational estimates. The longwave flux feedback is uniformly positive across the 3 ensembles and at least a factor of 2 larger than the shortwave flux feedback. This results in a positive net radiative flux feedback across the models with an average of 1.9±0.9 W m^{-2} K^{-1} and an average net surface flux feedback that is not statistically different from zero at the 95% level. Therefore, even though both the air-sea datasets and the models produce an increase in cloudiness as cold-tongue SSTs increase, all three models produce significantly weaker

306 net surface flux feedbacks than the observational estimates due to weaker negative
307 shortwave flux feedbacks and positive longwave flux feedbacks.

308 ## 4. Summary and Discussion

309 In this study we used yearly-initialized decadal hindcasts with 10 ensemble members for
310 each start date from four CMIP5 coupled climate models to assess simulations of 1970-
311 2009 equatorial Pacific SST, zonal wind stress, and surface flux trends. SSTs and zonal
312 wind stress were assessed against three data assimilations and four SST reconstructions,
313 and surface fluxes were assessed against CORE.2, OAFlux, and NOCv2 air-sea flux
314 datasets.

315 All four ensembles forecast a statistically significant warm-pool SST warming trend that
316 is within the observed estimates and a statistically significant cold-tongue SST warming
317 trend that is an order of magnitude larger than an ENSO residual estimated using SST
318 instrumental reconstructions. All forecasts produce a close relationship between the near-
319 equatorial zonal SST gradient across the equatorial Pacific and the zonal mean zonal
320 wind stress, indicating that, even if both the warm-pool and cold-tongue regions warm,
321 the strength of the Walker cell is a function of the *relative* warming between the warm-
322 pool and cold-tongue regions. Multi-model ensemble means formed using forecasts 6-10
323 years from initialization with 40 ensemble members do not produce an unambiguous
324 zonal SST gradient response to an increase in external forcing.

325 In the warm-pool region all year-1 forecasts produce SST trends similar to observations
326 but ocean mixed-layer and net surface heat flux trends with opposite sign to CORE.2,

OAFlux, and NOCv2 estimates. In addition, year-1 forecasts produce a positive shortwave feedback on decadal time scales, while 6-10 year forecasts produce a negative or statistically insignificant shortwave flux feedback on decadal time scales, suggesting a sensitivity to circulations forced by the initialized ocean state.

In the cold-tongue region all three initialized ensembles forecast a positive net radiative flux trend even though the shortwave flux trend is negative, i.e., for increasing cloudiness. This is inconsistent with the air-sea datasets, which uniformly show that the net surface radiative flux feedback is a damping of the underlying SSTs.

In this study we focus on near-equatorial fields at the air-sea interface and therefore cannot determine the role of large-scale circulation and the vertical structure of clouds and water vapor in the SST, zonal wind stress, and surface flux trends. Clearly, a positive net surface flux feedback in the cold-tongue region is contributing to the systematic positive SST trend in the cold-tongue region, but to what extent this is due to changes in the atmospheric circulation and structure of the clouds needs to be determined.

In a follow-up study we will investigate the processes that contribute to the positive longwave flux feedback and what causes the longwave flux feedback to be larger than the negative shortwave flux feedback in the cold-tongue region in year 6-10 forecasts. For example, a question that will be addressed is, can the positive longwave flux feedback be understood in terms of relatively constant high cloud temperatures due to rising cloud tops (the fixed anvil temperature hypothesis of Hartmann and Larson (2002) and discussed relative to CMIP3 models in Zelinka and Hartmann (2010))? What role do the

observed north-south asymmetries in cloud cover changes (discussed in Bellomo et al. (2014)) play in setting up the atmospheric circulations that impact equatorial cloud structure and water vapor feedbacks (e.g., Sun et al. 2006; Sun et al. 2009)? These investigations will advance the understanding of coupled feedbacks in the equatorial Pacific that determine how the cold-tongue region will respond to an increase in external forcing.

Acknowledgements: Matt Newman, Clara Deser, Kris Karnauskas, and an anonymous reviewer are acknowledged for feedback on this study. The World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and the climate modeling groups listed in this paper are acknowledged for producing and making their model output publicly available. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and leads development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. This work has been supported by grants from the NOAA OAR CVP program and NSF AGS 1125561.

References

Arora, V., and Coauthors, 2011: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. *Geophys. Res. Lett.*, **38**, L05805, doi:10.1029/2010GL046270.

Balmaseda, M.A., K. Morgensen, and A. Weaver, 2013: Evaluation of the ECMWF Ocean Reanalysis ORAS4. *Q. J. Roy. Met*., doi:10.1002/qj.2063.

Bellomo, K., A.C. Clement, J.R. Norris, and B.J. Soden, 2014: Observational and Model Estimates of Cloud Amount Feedback over the Indian and Pacific Oceans. *J. Climate*, **27**, 925–940. doi: http://dx.doi.org/10.1175/JCLI-D-13-00165.1.

Bendat, J.S., and A.G. Piersol, 2000: Random data analysis and measurement procedures. ISBN 0471317330, P113.

Berry, D.I., and E.C. Kent, 2011: Air–sea fluxes from ICOADS: the construction of a new gridded dataset with uncertainty estimates. *Int. J. Climatol*., **31**, 987–1001. doi: 10.1002/joc.2059.

Branstator, G., and H. Teng, 2010: Two limits of initial-value decadal predictability in a CGCM. *J. Climate*, **23**, 6292–6311. doi:10.1175/2010JCLI3678.1.

Bunge, L., and A.J. Clarke, 2009: A verified estimation of the El Nino index NINO3.4 since 1877. *J. Climate*, **22**, 3979-3992.

381 Cane, M.M., A.C. Clement, A. Kaplan, Y. Kushnir, D. Pozdnyakov, R. Seager, S.E.
382 Zebiak, and R. Murtugudde, 1997: Twentieth-century sea surface temperature
383 trends. *Science*, **275**, 957–960.

384 Carton, J.A., and B.S. Giese, 2008: A reanalysis of ocean climate using Simple Ocean
385 Data Assimilation (SODA). *Mon. Wea. Rev.,* **136**, 2999-3017.

386 Chang, Y.-S., S. Zhang, A. Rosati, T. Delworth, and W.F. Stern, 2013: An assessment of
387 oceanic variability for 1960-2010 from the GFDL ensemble coupled data
388 assimilation, *Clim. Dyn*., **40**, 775-803, doi:10.1007/s00382-012-1412-2.

389 Clarke, A.J., and A. Lebedev, 1996: Interannual and decadal changes in equatorial wind
390 stress in the Atlantic, Indian, and Pacific Oceans and the Eastern Ocean coastal
391 response. *J. Climate*, **10**, 1722–1729.

392 Compo, G.P., and P.D. Sardeshmukh, 2010: Removing ENSO-related variations from the
393 climate record. *J. Climate*, **23**, 1957-1978.

394 Delworth, T.L., and Coauthors, 2006: GFDL's CM2 global coupled climate models. part
395 I: Formulation and simulation characteristics. *J. Climate*, **19**, 643–674.

396 Deser, C., A.S. Phillips, and M.A. Alexander, 2010: Twentieth century tropical sea
397 surface temperature trends revisited. *Geophys. Res. Lett*., **37**, L10701.

398 Gordon, C., C. Cooper, C.A. Senior, H. Banks, J.M. Gregory, T.C. Johns, J.F.B. Mitchell,
399 and R.A. Wood, 2000: The simulation of SST, sea ice extents and ocean heat

transport in a version of the Hadley Centre coupled model without flux adjustments. *Clim. Dyn*., **16**, 147–168.

Hartmann, D.L., and K. Larson, 2002: An important constraint on tropical cloud-climate feedback. *Geophys. Res. Lett*., **29**(20), 1951, doi:10.1029/2002GL015835.

Ishii, M., A., Shouji, S. Sugimoto, and T. Matsumoto, 2005: Objective analyses of SST and marine meteorological variables for the 20th century using ICOADS and the Kobe Collection. *Int. J. Climatol.,* **25**, 865–879.

Kaplan, A., and Coauthors, 1998: Analyses of global sea surface temperature 1856–1991. *J. Geophys. Res*., **103**, 18,567–18,589, http://dx.doi.org/10.1029/97JC01736.

Karnauskas, K.B., R. Seager, A. Kaplan, Y. Kushnir, and M.A. Cane, 2009: Observed strengthening of the zonal sea surface temperature gradient across the equatorial Pacific Ocean. *J. Climate*, **22**, 4316-4321.

Kim, H.M., P.J. Webster, and J.A. Curry, 2012: Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts. *Geophys. Res. Lett.*, **39**, L10701, doi:10.1029/2012GL051644

L'Heureux, M., S. Lee, and B. Lyon, 2013: Recent multidecadal strengthening of the Walker circulation across the tropical Pacific. *Nature Clim. Change,* **3**, 571–576.

Meng, Q., and Coauthors, 2012: Twentieth century walker circulation change: Data analysis and model experiments. *Clim. Dyn.* **38**, 1757–1773.

419 Merryfield, W.J., W. Lee, G.J. Boer, V.V. Kharin, J.F. Scinocca, G.F. Flato, R.S.
420 Ajayamohan, Y. Tang, and S. Polavarapu, 2013: The Canadian Seasonal to
421 Interannual 1 Prediction System. Part I: Models and initialization. *Mon. Wea. Rev.*,
422 doi:10.1175/MWR-D-12-00216.1.

423 Norris, J.R., 2005: Trends in upper-level cloud cover and surface divergence over the
424 tropical Indo-Pacific Ocean between 1952 and 1997. *J. Geophys. Res*., **110**,
425 D21110, doi:10.1029/2005JD006183.

426 Power, S.B., and I.N. Smith, 2007: Weakening of the Walker Circulation and apparent
427 dominance of El Niño both reach record levels, but has ENSO really changed?
428 *Geophys Res Lett*. **34**, L18702, doi:10.1029/2007GL030854.

429 Rayner, N.A., and Coauthors, 2003: Global analyses of sea surface temperature, sea ice,
430 and night marine air temperature since the late nineteenth century. *J. Geophys.*
431 *Res*., **108**, 4407, doi:10.1029/2002JD002670.

432 Smith, T.M., R.W. Reynolds, T.C. Peterson, and J. Lawrimore, 2008: Improvements to
433 NOAA's historical merged land-ocean surface temperature analysis (1880-2006).
434 *J. Climate,* **21**, 2283-2296.

435 Solomon, A., and M. Newman, 2011: Decadal predictability of tropical Indo-Pacific
436 Ocean temperature trends due to anthropogenic forcing. *Geophys. Res. Lett.*, **38**,
437 L02703, doi:10.1029/2010GL045978.

438 Solomon, A., and M. Newman, 2012: Reconciling disparate 20th century Indo-Pacific
439 ocean temperature trends in the instrumental record. *Nature Climate Change*, **2**,
440 691-699, doi:10.1038/NCLIMATE1591.

441 Sun, D.-Z., T. Zhang, C. Covey, S. Klein, W.D. Collins, J.J. Hack, J.T. Kiehl, G.A.
442 Meehl, I.M. Held, and M. Suarez, 2006: Radiative and Dynamical Feedbacks
443 Over the Equatorial Cold-tongue: Results from Nine Atmospheric GCMs. *J.*
444 *Climate*, **19** , 4059-4074.

445 Sun, D.-Z., Y. Yu, and T. Zhang, 2009: Tropical Water Vapor and Cloud Feedbacks in
446 Climate Models: A Further Assessment Using Coupled Simulations. *J. Climate*,
447 **22**, 1287-1304.

448 Taylor, K.E., R.J. Stouffer, and G.A. Meehl, 2012: An Overview of CMIP5 and the
449 Experiment Design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498. doi:10.1175/BAMS-D-
450 11-00094.1.

451 Tokinaga, H., S.-P. Xie, A. Timmermann, S. McGregor, T. Ogata, H. Kubota, and Y.M.
452 Okumura, 2012: Regional Patterns of Tropical Indo-Pacific Climate Change:
453 Evidence of the Walker Circulation Weakening. *J. Climate*, **25**, 1689–1710.

454 Trenberth, K.E., 2002: Changes in tropical clouds and radiation, *Science*, **296**, 2095a.

455 Vecchi, G.A., and B.J. Soden, 2007: Global warming and the weakening of the tropical
456 circulation. J. *Climate*, **20**, 4316-4340.

457 Vecchi, G.A., B.J. Soden, A.T. Wittenberg, I.M. Held, A. Leetmaa, and M.J. Harrison,
458 2006: Weakening of tropical Pacific atmospheric circulation due to anthropogenic
459 forcing. *Nature*, **441**, 73-76.

460 Wielicki, B.A., and Coauthors, 2002a: Evidence for large decadal variability in the
461 tropical mean radiative energy budget. *Science*, **295**, 841-844.

462 Wielicki, B.A., and Coauthors, 2002b: Response. *Science*, **296**, 2095a.

463 Woodruff, S. D., H. F. Diaz, J. D. Elms, and S. J. Worley, 1998: COADS release 2 data
464 and metadata enhancements for improvements of marine surface flux fields. *Phys.*
465 *Chem. Earth*, **23**, 517–526.

466 Worley, S. J., S. D. Woodruff, R. W. Reynolds, S. J. Lubker, and N. Lott, 2005: ICOADS
467 release 2.1 data and products. *Int. J. Climatol*., **25**, 823–842.

468 Yeager, S.G., and W.G. Large. 2008. *CORE.2 Global Air-Sea Flux Dataset*. Research
469 Data Archive at the National Center for Atmospheric Research, Computational
470 and Information Systems Laboratory. http://dx.doi.org/10.5065/D6WH2N0S.
471 Accessed 1 Jan 2013.

472 Yu, L., and R.A. Weller, 2007: Objectively Analyzed air-sea heat Fluxes for the global
473 ice-free oceans (1981–2005). *Bull. Ameri. Meteor. Soc*., **88**, 527–539.

474 Zelinka, M.D., and D.L. Hartmann, 2010: Why is longwave cloud feedback positive? *J.*
475 *Geophys. Res*., **115**, D16117, doi:10.1029/2010JD013817.

Tables

(W m^{-2} K^{-1})	CORE.2	OAFLUX	NOCv2	MME
SW	**-4.1**	**-6.5**	**-6.2**	**-5.6±1.3**
LW	**-2.4**	*-1.0*	**2.5**	*-0.3±2.5*
SW+LW	**-6.6**	**-7.4**	**-3.7**	**-5.9±1.9**
NET	**-12.0**	**-13.7**	**-13.5**	**-13.1±0.9**

Table 1: 1983-2006 cold-tongue annual mean observational estimates of surface flux feedback coefficients, in units of W m^{-2} K^{-1}. Right column shows the mean over the three air-sea flux datasets +/- one standard deviation. Statistically significant (insignificant) coefficients at the 95% level based on a two-tailed Student's *t*-test highlighted with bold (italics) font.

(W m^{-2} K^{-1})	CanCM4 YR6-10	GFDL YR6-10	HadCM3 YR6-10	MME
SW	*-0.5*	**-1.5**	**-1.7**	**-1.2±0.7**
LW	**1.9**	**2.9**	**4.6**	**3.1±1.4**
SW+LW	**1.4**	*1.4*	**2.9**	**1.9±0.9**
NET	**1.4**	*-0.1*	*1.1*	*0.8±0.8*

Table 2: 1970-2009 cold-tongue annual mean year 6-10 forecast surface flux feedback coefficients, in units of W m^{-2} K^{-1}. Right column shows the mean over the three ensembles +/- one standard deviation. Statistically significant (insignificant) coefficients at the 95% level based on a two-tailed Student's *t*-test highlighted with bold (italics) font.

Table Captions:

Table 1: 1983-2006 cold-tongue annual mean observational estimates of surface flux feedback coefficients, in units of W m^{-2} K^{-1}. Right column shows the mean over the three air-sea flux datasets +/- one standard deviation. Statistically significant (insignificant) coefficients at the 95% level based on a two-tailed Student's *t*-test highlighted with bold (italics) font.

Table 2: 1970-2009 cold-tongue annual mean year 6-10 forecast surface flux feedback coefficients, in units of W m^{-2} K^{-1}. Right column shows the mean over the three ensembles +/- one standard deviation. Statistically significant (insignificant) coefficients at the 95% level based on a two-tailed Student's *t*-test highlighted with bold (italics) font.

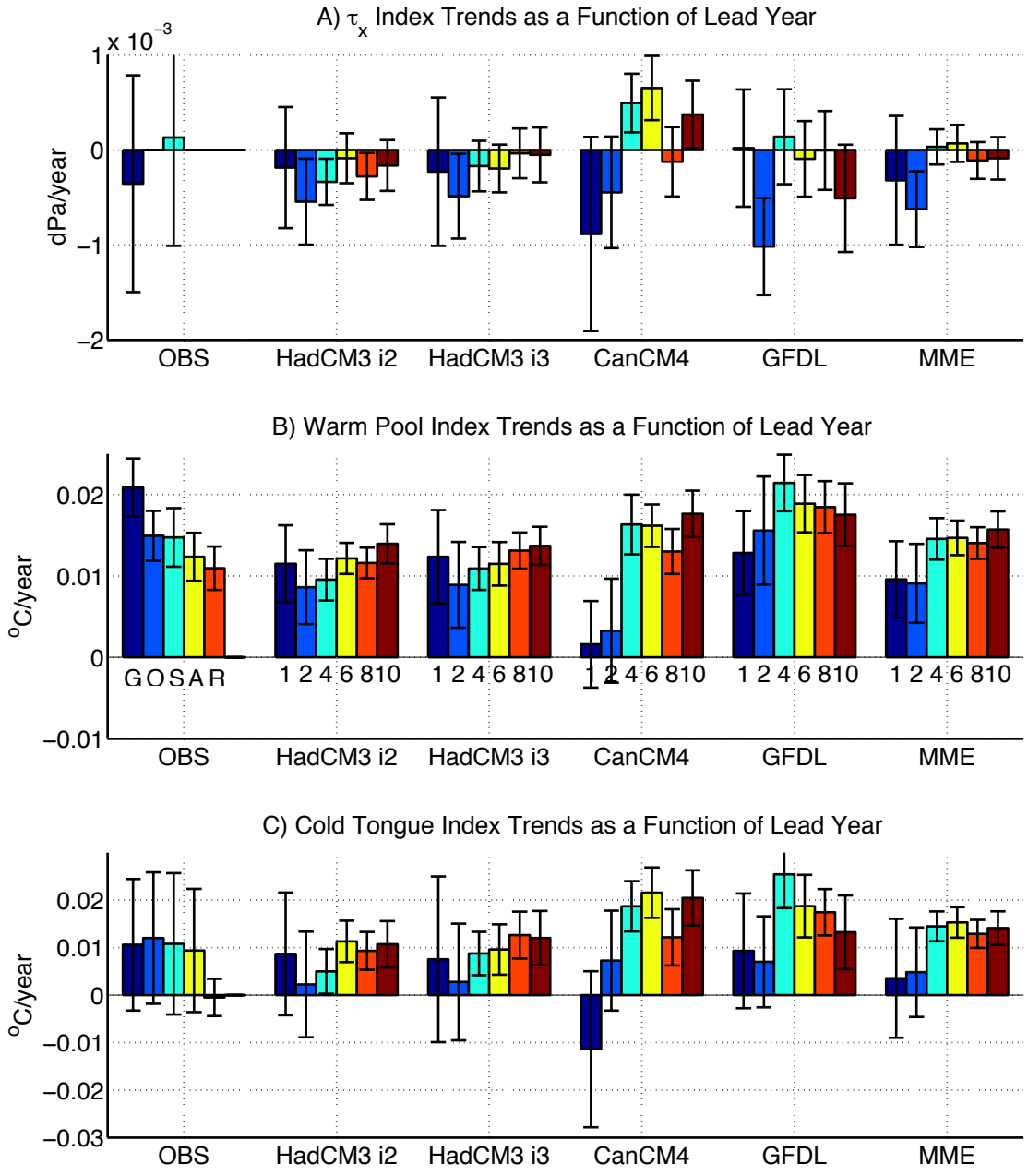
Figure Captions:

Figure 1: A) τ_x index 1970-2009 linear trends as a function of lead year, in units of dPa year^{-1}. B) Warm-pool index 1970-2009 linear trends as a function of lead year, in units of °C year^{-1}. C) Cold-tongue index 1970-2009 linear trends as a function of lead year, in units of °C year^{-1}. Trends for lead years 1,2,4,6,8,10 shown for ensemble mean hindcasts, as indicated in (B). MME indicates multi-model ensemble mean. Observed 1970-2009 trends estimated using three data assimilations (marked, "G","O","S") and a best estimate from 4 SST reconstructions with (marked "A") and without (marked "R") ENSO variability. Error bars show 95% confidence intervals.

Figure 2: Scatterplot of 1970-2009 warm-pool minus cold-tongue SST trends vs τ_x trends for observations and decadal hindcasts shown in Fig. 1 of the manuscript, in units of °C year^{-1} and dPa year^{-1}, respectively. Observations are shown with large green filled circles. Ensemble mean year 1 forecasts shown with black filled circles. Ensemble mean year 2-5 forecasts shown with blue filled circles. Ensemble mean year 6-10 forecasts shown with red filled circles. Multi-model ensemble means shown with larger filled circles with yellow centers. Lines indicate confidence intervals estimated with a two-sided P=0.05 level. The thin black dash diagonal line indicates a τ_x trend of -1.0 dPa year^{-1} for a SST gradient trend of 5.e-3 °C year^{-1}.

Figure 3: 1983-2006 annual mean equatorial Pacific surface energy flux trends, in units of W m^{-2} year^{-1}. Left) Warm-pool. Right) Cold-tongue. A) CORE.2. (black error bars), OAFlux (red bars), NOCv2 (green bars). B) HadCM3-i2. C) CanCM4. D) GFDL CM2.1. NET is equal to SW+LW+SH+LH. Error bars show 95% confidence intervals. Year-1

518 forecast of trends shown with black error bars. Year 6-10 forecast shown with red error
519 bars. All fluxes are positive downward. Note larger range of fluxes in (A).



520

521 **Figure 1:** A) τ_x index 1970-2009 linear trends as a function of lead year, in units of
522 dPa/year. B) W̄arm-pool index 1970-2009 linear trends as a function of lead year, in units
523 of °C/year. C) Cold-tongue index 1970-2009 linear trends as a function of lead year, in
524 units of °C/year. Trends for lead years 1,2,4,6,8,10 shown for ensemble mean hindcasts,
525 as indicated in (B). MME indicates multi-model ensemble mean. Observed 1970-2009
526 trends estimated using three data assimilations (marked, "G","O","S") and a best estimate
527 from 4 SST reconstructions with (marked "A") and without (marked "R") ENSO
528 variability. Error bars show 95% confidence intervals.

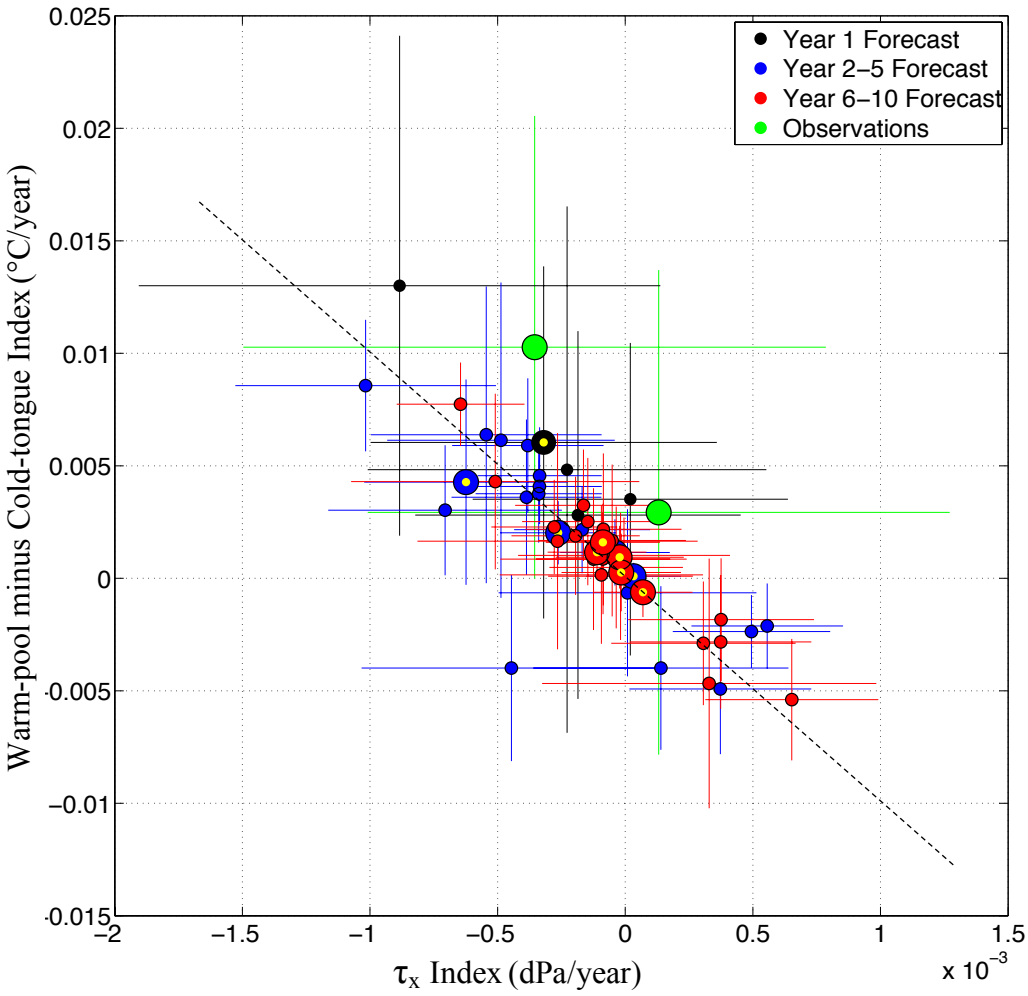


Figure 2: Scatterplot of 1970-2009 warm-pool minus cold-tongue SST trends vs τ_x trends for observations and decadal hindcasts shown in Fig. 1 of the manuscript, in units of °C/year and dPa/year, respectively. Observations are shown with large green filled circles. Ensemble mean year 1 forecasts shown with black filled circles. Ensemble mean year 2-5 forecasts shown with blue filled circles. Ensemble mean year 6-10 forecasts shown with red filled circles. Multi-model ensemble means shown with larger filled circles with yellow centers. Lines indicate confidence intervals estimated with a two-sided P=0.05 level. The thin black dash diagonal line indicates a τ_x trend of -1.0 dPa year^{-1} for a SST gradient trend of 5.e-3 °C year^{-1}.

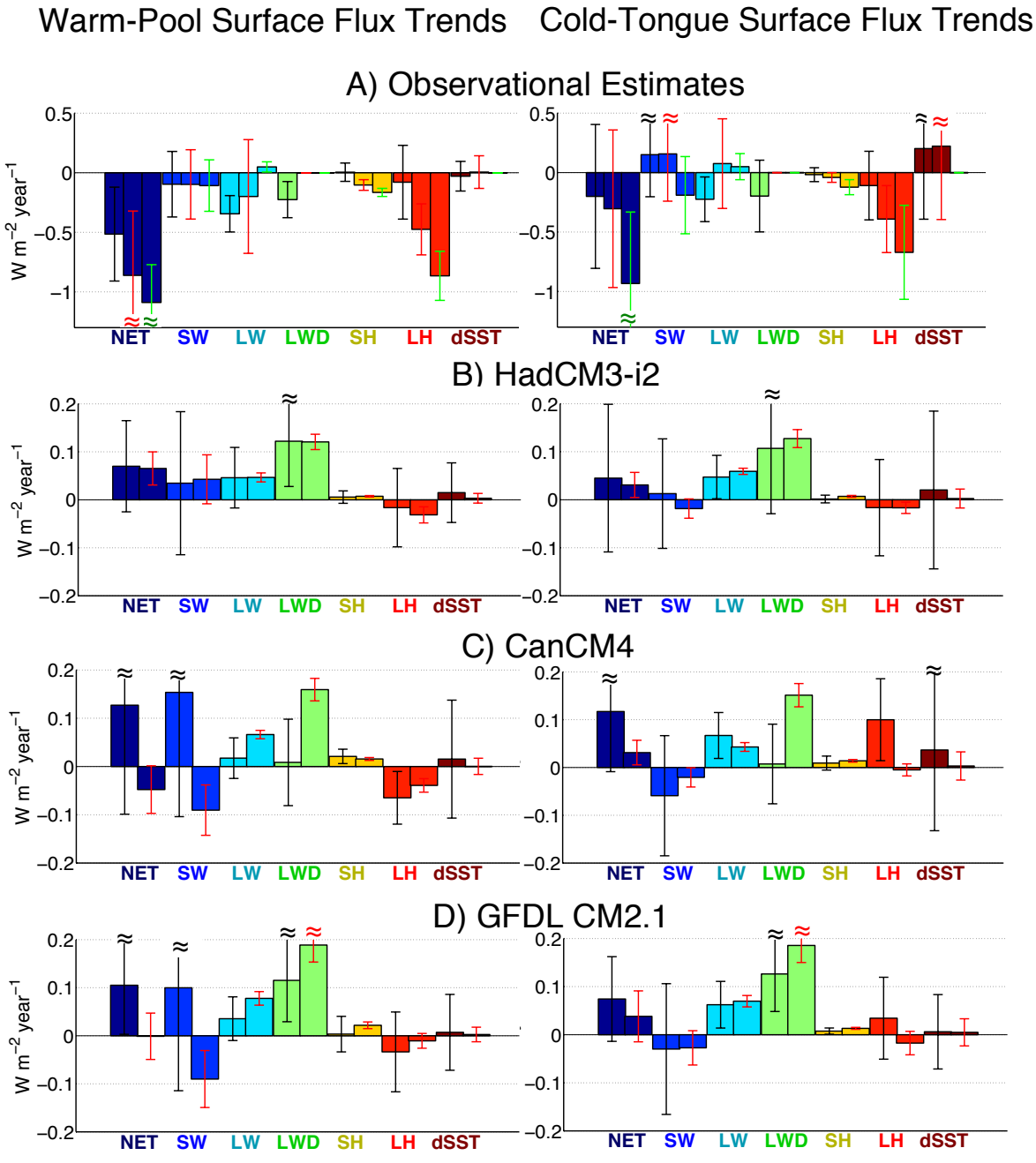


Figure 3: 1983-2006 annual mean equatorial Pacific surface energy flux trends, in units of W m^{-2} $year^{-1}$. Left) Warm-pool. Right) Cold-tongue. A) CORE.2. (black error bars), OAFlux (red bars), NOCv2 (green bars). B) HadCM3-i2. C) CanCM4. D) GFDL CM2.1. NET is equal to SW+LW+SH+LH. Error bars show 95% confidence intervals. Year-1 forecast of trends shown with black error bars. Year 6-10 forecast shown with red error bars. All fluxes are positive downward. Note larger range of fluxes in (A).